[CCRMA DSP Seminar, Thursday June 3, 3:15pm, CCRMA Ballroom]

A Robust Constant Q Spectrum for Polyphonic Pitch Tracking

Benjamin Blankertz

Institute for Mathematical Logic
University of Münster, Germany
blanker@uni-muenster.de

A straight forward approach to polyphonic pitch tracking and related tasks is based on a constant Q spectral transform (Q denotes the frequency to resolution ratio). As shown in Brown 1991 [1] such a transform can be implemented as a bank of Fourier filters of variable window width. The geometrical spacing of the center frequencies makes the components of the constant Q spectrum directly relate to musical frequencies of the chromatic scale (in equal temperament). Front-end processing with a constant Q transform makes problems like note identification and instrument classification (simplifying matters) available to pattern recognition methods, such as representing the input signal's spectrum as NNLS (nonnegative least squares) solution in a database of single tone spectra.

Besides from many advantages two problems remain in this approach.
1. Owing to spectral leakage the magnitude of a constant Q transform (as well as the Fourier transform) is sensitive to phase changes in the stimulus, a phenomenon also termed 'Fourier uncertainty'. For employing pattern recognition methods a spectral representaion is desirable that is invariant under phase changes.
2. We have the usual dilemma of trading time against frequency resolution. Time resolution is a fixed magnitude, namely the reciprocal value of the length of each analysis window, whereas frequency resolution cannot be pinpointed in such a definite way. In a bank of filters frequency resolution is given by the bandwidth of each filter. But for real world signals the accuracy of frequency estimates is strongly affected by noise and interference between different sinusoidal contributions of neighboring frequencies. This poses the problem of getting the best possible frequency determination at a given (high) time resolution.

'A high resolution fundamental frequency determination based on phase changes of the Fourier transform' (Brown/Puckette 1993 [2]) is the frequency determination method of the phase vocoder. This technique makes use of the phase information in the Fourier spectrum, but nevertheless systematic errors depending on the phase of the input signal remain. A detailed analysis of this effect in the case of pure sinusoidal waves yields an explicit form of the error function in terms of frequency and phase of the stimulus. This structural information can be used in different ways for increasing the accuracy of the frequency determination. In this talk a method of adaptive windows is presented in which each peak of an initially calculated Fourier spectrum is analyzed seperately. From a preestimate of the corresponding frequency the parameters of a 3-term Blackman window are determined which minimize the expected error with respect to the preestimated frequency. There is quasi no error in this estimation (maximal error far below 1 milli cent) for pure sine waves in a broad frequency range. Even in cases where less than half a period of a sinusoidal falls in the analysis window precise results are achieved. Though having mathematically a quite involved derivation the generalization to complex tones shows very promising results.

Analyzing a piece of music frame-by-frame in this way yields, for each frame, a list of peaks of spectral intensity. Each entry of the list consists of the amplitude and a high-resolution estimate of the frequency. The proposed robust constant Q spectrum is the abstraction of such a peak list to a vector whose components correspond to fixed frequencies that are equally spaced in the log frequency domain. This vector looks formally like the usual constant Q transform but it has two advantages over the latter. (1) All non-zero values in that vector steam from actual peaks in spectral intensity and (2) the frequency information is not phase sensitive and it is reliable even at a high time resolution, i.e. the two problems discussed above are resolved.

[1] Judith C. Brown (1991): *Calculation of a constant Q spectral transform*, Journal of the Acoustical Society of America 89 (1), 425–434.

[2] Judith C. Brown, Miller S. Puckette (1993): *A high resolution fundamental frequency determination based on phase changes of the Fourier transform*, Journal of the Acoustical Society of America 93 (2), 662–667.